RESEARCH ARTICLE

Data Discovery Tools and Marketing-Comparative Case Study

Tuţă M*, Zară IA, Munthiu MC

Faculty of Marketing, Academy of Economic Studies, Buchares, Romania.

Corresponding Author: E-mail*: *mihaela.tuta@gmail.com

Abstract

During last decades, technologies developed very much and the existing software applications help companies to become more competitive. The data about customers, which are collected daily, can be easily transformed into information using the correct tools. Business Intelligence tools main objective is to offer the business users greater support for decisions. The current research article begins with a short presentation of what data mining is and how it can be used in the marketing field. It continues with a short presentation on how Business Intelligence tools may integrate Data Mining algorythms and are able to offer answers to some marketing questions. The last part of the current studyrepresents a comparison among the first three data discovery tools existing on the market.The data discovery tools are a subdivision on the Business iNtelligence Tools, providing the users the possibility of building ad-hoc queries and finding the answers to spntaneous questions.

Keywords: *Ad-hoc queries, Business Intellingence, Data mining, Data discovery tools, Marketing.*

Introduction

Having access to the right information at the right time is essential in today's battle between companies to win their customer's respect and loyalty. To do this, it is no longer sufficient to offer very good products and services to customers, but to build a long time relationship with them. For small, niche business, this is not very hard, because the employees may know the consumers directly, building a long time relationship and getting to know them them very well. Knowing the customers better (e.g: which are their shopping habits, change in their interests and behaviour) might be a challenge for bigger multinational organizations [1].The good news for this kind of organization is that on the IT market there are a lot of solutions that can help companies achieving their goal of understanding their customers better. During the last decades, the IT technologies developed a lot: the hardware is cheaper, the data processing technologies are being developed constantly (e.g.: cloud computing, parallel processing) and networks become more secure (WMN) [2]and more accessible for the public).Even if data may be stored in different databases inside an organization, they can be collected in a common data warehouse that becomes the memory of the organization and provide the raw material for the data mining process. Data Mining is a process of extracting information from a large data volume, without

prior assumption or model.The cureent study shows how data mining algorythms can be used to offer the answers to different marketing questions. Usually, Business Intelligence tools integrate Data Mining algorythms to enable the users to obtain the best results. Business Intelligence Tools (BI) are classified into traditional BI Reporting and Data Discovery tools that were developed durintg the last three years. When using the Data Discovery tools, the users can build their own graphs based on their informational needs and experience.The current study contains a comparison between the best well known Data Discovery tools existing on the market, using, mainly different data available on the existing demos.

Data Mining and Marketing

Data mining main objective is to extract knowledge from large amounts of data using sophisticated models and techniques. What is special about the datamining is that there is no assumption how the result may look like.Data mining can help organizations to get customers' insights and build an effective CRM strategy, building personalized interaction. Datamining can be used in all the phases of the customer lifecycle [4].

- Getting new customers by using different aquisition models to identify potentially profitable customers;
- Establish and develop the current base of existing customers by using cross-/deep-/up-selling models that can reveal the current potential of purchasing of the current customers;
- Retain customers by using attrition models that help to identify the customers who are more likely to leave.

Data Mining and Business Intelligence

Companies have been using data mining for a long time, and recent advances in technology helped adopting some datamining processes on a larger scale. Companies can analyze increasing amounts of data in order to reduce the costs and increase the profitability [4] . „Business Intelligence (BI) represents the capability to look inside a business and the environment in which it operates to fundament the most productive and profitable Decisions" [5-6].

In 1989, Howard Dresner, researcher at Gartner Group defined Business Intelligence as "a set of concepts and methods for improving the decision process, using support systems based on facts".Data Mining algorythms are integrated into Business Intelligene tools, which are mainly used to support different type of decision.Firstly, the main objectives of using a Business Intelligence tool were: to increase the organizations' income, to understand the customers better and to identify the customers who have a higher lifetime value. Business Intelligence can succesfully be used in different areas such as: market analysis, market segmentation, marketing campaign management, distribution channels management, customer relationship management, sales force management [7].

BI and Marketing Analysis

the main answers marketing analysis tries to answer are: „who buys the goods produced by a company?" „how do customers react to the price change?", „which products or services generate what percentage of revenues?", „what are the sales trends for each product or sevice?". For companies who daily collect a lot of information about their customers, a well configured Business Intelligence system may answer the above questions and the market analysis may be done in almost real time and can offer a good view on the current and future revenue trends. Having the right information the companies can choose the right policies to retain the customers, gain new customers and enlarge their portofolio. One of the most profitable uses of BI is to better understand the relationship among customers, products, services and revenue generation.

Customer Segmentation

In the past customer segmentation was done using demographic or geograhic segmentation. In the business to business area, the main criteria used in customer segmentation were: the industry, the role of the organization in the distribution chanel and the profit. With the help of data mining, and combining the results with the above mentioned criteria, the companies can define narrower customer segments , understand better the needs and values of the customers, create products and services to better respond to their demands and develop ways to reach and acquire new customers or expand the business with the current customers.

CRM (Customer Relationship Management)

CRM have different meanings for different people. To those who sell these type of solutions, CRM is a software package, which has both transational and BI functionality. To those who develop BI systems, CRM means a packaged software applications used to analyse the consumer behaviour and sales force performance. To the comapnies who want to improve the level of their revenue, CRM should have transactional functionalities and BI and reporting functionalities. Both type of functionalities should use a common database where information about customers are stored and can be used in the selling activities. The CRM systems are usually used to automate the marketing campaigns and can be enhanced with BI functionalities. If the results delivered by this type of packaged software are associated with changes in specific marketing or sales processes, then companies can increase their venues.

Existing BI Tools

According to Jeremy Kolb, there are two major type of BI tools: traditional BI reporting tools and data discovery tools. Traditional BI reporting tools main objective is to offer a view on what is happening inside a business and report on those findings.BI tools offer the posibility to focus on a certain area of the business (e.g: sales, finance, employee productivity) and provides a detailed analysis for all the factors that influence that area. Reports generated with a traditional BI are tailored and are generated only when a request is made. Another element included in the traditional BI tools is the Dashboard, which is real-time generated and offers the possibility to monitor the

KPIs on a daily basis. The Dashboards provide an overview of large segments.

The new trend in BI is Data Discovery and predictive analysis, which developed starting back in 2010. The main aim of data discovery is to provide a good user experience, trying to answer their emerging questions in an interactive and graphical environment. Users can build their own graphs, based on their experience and depending on their informational needs. The well known functionalities of a Data Discovery BI tool are Data Visualization and Ad-Hoc queries [4].

Ad-Hoc queries being built spontaneously by the user, without being developed in advance, compared with the Dashboards from traditional BI tools. Ad-Hoc queries are versatile and enable users to find answers to their non- standard questions.

Data Visualization provide context, history and future projections to data in a simplistic and interactive way. It allows the users to explore data, looking at what they think it is important. The major advantages offered by Data Discovery tools are [8].

Universal adaptability- by using a single analytics and data discover platform anyone can make some insightful decisions;

Visibility into the unknown- users have the opportunity to discover unexpected insights hidden in the data. Those insights can help them identifying strategies, business opportunities and threats.

Self- Service Discovery- users can explore data to any level of detail, wothout being dependent on IT departmment, accelerating the process of decision making.

Columns	Customer Segment	⊞ YEAR(Order Date)
Rows	Order Priority	

| | Customer Segment / Order Date | | | | | | | |
| | Consumer | | Corporate | | Home Office | | Small Business | |
Order Priority	2011	2012	2011	2012	2011	2012	2011	2012
Critical	$11,895	$340	$23,001	$36,682	$11,617	$642	$14,004	$10,103
High	$27,359	$25,411	$14,607	$31,079	$25,014	$35,002	$47,103	$34,005
Medium	$9,789	$18,452	$20,499	$26,537	$8,266	$12,651	$12,251	$6,555
Low	$6,814	$8,730	$42,810	$34,713	$6,616	$15,556	$27,844	$12,819
Not Specified	$14,039	($3,253)	$30,532	$24,850	$29,962	$7,057	($2,568)	$3,971

Fastest to actionable insights- users can fastly turn an insight into action because thay can discover more quickly what hides inside the data.

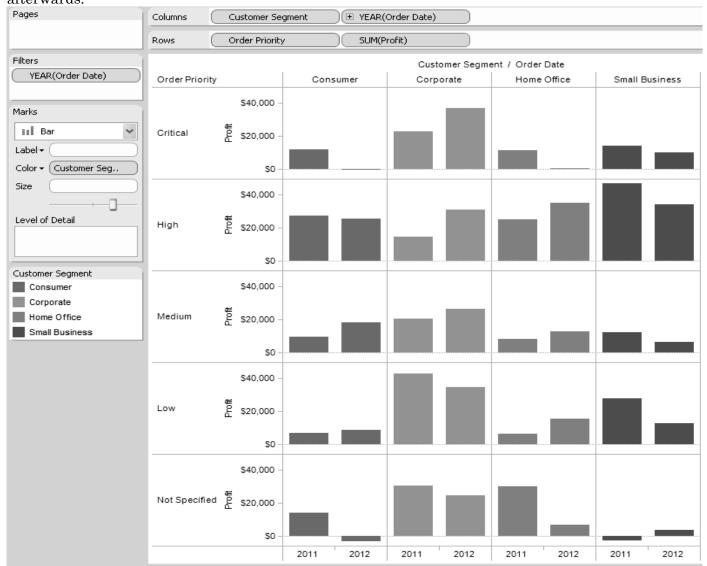## Data Visualization & Data Discovery Tools-Comparative Study

There are several tools providing Data Discovery functionalities. According to Applied Data Labs, the most popular Data Discovery tools are: Tableau, QlikView, Tibco Spotfire and Microsoft Excel [9] . In the next section we will present these tools and will provide a short comparison between them with the help of few test cases.

Tableau software is a desktop-based tool that allows user to use the mouse to drag and drop pieces of data onto canvases where they can display the data in different types of charts, graphs, and maps. It connects directly to different data sources (local and remotely) [9]. Connection to a server provides a large list of alternatives, from Claudera Hadoop, IBM DB2, Microsoft SQL

Server to Oracle and Sybase. Tableau offers the users the chance to publish views and dashboards to the Web, and has a mobile client The mobile functionalities are ranked weaker than the ones of the competitors because it offers limited flexibility and all the development should be done prior, on a desktop client [9]. Using the desktop client is very intuitive and user friendly. No technical background is required for the person using Tableau desktop. The user should know know very well what he wants to achieve using this tool.For example, we used the available sample-Superstore Sales (excel) from Tableau. We wanted to see what is the yearly profit brought by each customer segment depending on the order priority.

The analysis can go further and create some calculation fields to identify which is the most profitable customer segment.The drill up and drill down functionalities work well. The only

disatisfactor is that when turning the table view into graphical view, one dimension is lost in transition and it needs to be manually added, afterwards.

Pages				
Columns	Customer Segment	⊞ YEAR(Order Date)		
Rows	Order Priority	SUM(Profit)		

	Customer Segment / Order Date			
Order Priority	Consumer	Corporate	Home Office	Small Business

Filters
YEAR(Order Date)

Marks
▮▮ Bar ▾
Label ▾
Color ▾ Customer Seg..
Size
Level of Detail

Customer Segment
■ Consumer
■ Corporate
■ Home Office
■ Small Business

Tableau's data mining capabilities are limited--pattern recognition, correlation analysis, outlier identification--all of these are not provided by the tool itself. Unless user considers charts and graphs which help a human identify these valuable pieces of information, Tableau doesn't do much in this area. Qlikview is a BI software that combines various features such as: data analysis, data manipulation and dynamic presentation. It can be used for reporting, forecast and general data analysis. Qlik can be easily used in any field where real-time data analysis is required [10].

QlikView has many of the same drag-and-drop user interface components that Tableau does, and they appear to be very similar. QlikView is often considered the leader in interactive visualizations, and offers great flexibility for end

users who want to manipulate their view of the data. QlikView has very heavy deployments which require technical resources to develop the bones of the deployment and connect it to data

sources before end users can even get their hands on it [9]. QlikView requires the most development to be done on a desktop, and then the resulting analyses can be manipulated using mobile devices.We downloaded the Qlik View Personal edition full version trial from the official site, but couldn't use it too much due to the strong technical skills that are required to build and deploy QlikView. The version we downloaded is designed for the IT employees to load data, do the aggregations and calculations and customize the application that will be used by the business. The objective of this article is to present these products from the business point of view.

Data used by QlikView can be imported from various sources, databases, datawarehouses, excel spreadsheets. After creating an account, we logged in to the site and explored Qlik using a demo for the retail product performance. The demo was designed to demonstrate how QlikView can bring visibility to the performance of products sold in retail stores by letting the user to explore sales and product productivity, across any number of customers, items, or demographic dimensions, in order to better analyze and react to business trends [11].

The experience was positive because the environment is very user friendly. The first two sheets in the demo contain detailed information about the data used- Fine Foods stores, a grocery chain, which has stores located in three US regions: Northeast, South and West. The demo contains information by region, store, product category and product group from an overall perspective down to an individual product level. Users can build various analysis: store analysis, product analysis, product change analysis, customer profile analysis and a space simulation.

For example, retail Product Performance demo allows users to perform a product change analysis to determine if a change in how a product is bundled could increase sales of the product(s).
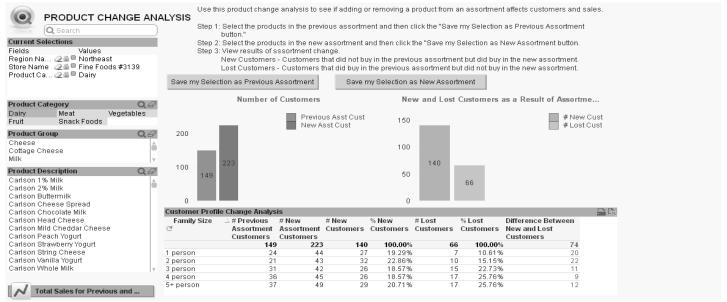

Fig. 3: The product change analysis

Source:http://eu.demo.qlikview.com/QvAJAXZfc/opendoc.htm?document=qvdocs/Retail%20Product%20Performance.qvw&host=Demo11
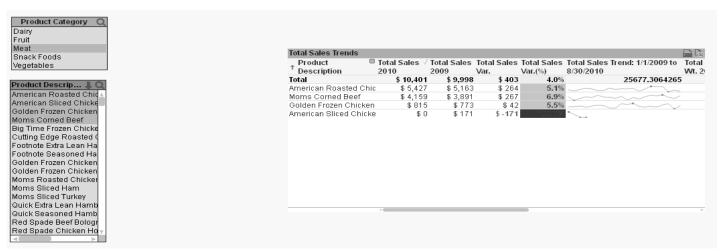

Fig. 3: The sales trends for product categories.

Source:http://eu.demo.qlikview.com/QvAJAXZfc/opendoc.htm?document=qvdocs/Retail%20Product%20Performance.qvw&host=Demo11

We built a test demo page where we wanted to see what are the sales trends for product categories. The custom reports are easy to built, using the same drag and rop functionalities, as Tableau.

The results shows like this:
Tibco's Spotfire is another strong competitor in the data discovery space. Spotfire offers capabilities in data visualization and exploration.

The costs to buy and deploy this solution are high. The developmant and deoployment should be done by technical experts. The end users have the posibility to use the reports using a mobile version or a desktop/laptop [9].

The data can be imported from various sources in the same way as for Tableau and QlikView.Once again, we used the demos available on the site. This time we chose the performance attrition demo. The views are more technical and not so detailed as the ones from Qlik. There is no posibility the user to built his own reports. The filters are easy to implement , to chose industry sectors and customers. What was better compared to the previous products is the fact that Spotfire gives the opportuniy to compare the customer's performance with industry performance using benchmarking.

Microsoft Excel is among the more capable data discovery tools on the market. Compared with previous products, Excel cannot be run from mobile devices and it is not very to use by the unexperienced users. Excel is very technical. For someone who knows how to use it, it is far and away the most flexible and capable of the data discovery tools. There is no limitationn in the questions users can ask. In a previous article we presented how Excel can be used with another CRM tool to discover new opportunities and tailor a service ofering for an important customer of a large IT company.

Results and Discussions

We decided to compare sum up the attributes we considered important fro the Data Discovery tool in the table below:

Table 1: Major attributes for the data discovery tools- comparison

	Importing Data	Prices to buy, deploy and implement	User Friendly	Mobile use	Ease of use
Tableau	Various Sources	Medium	High	Yes	High
QlikView	Various Sources	Medium	High	Yes	High
Spotfire	Various Sources	High	Medium	Yes	Medium
Excel	Various Sources	Included in the Office Licence	Low	No	Low

The main aim of the comparative study was to see what are the functionalities of the most important Data Discovery Tools from the market. Further studies can be conducted to see what is the impact of the decisions made based on the results return by such a tool.From a marketing perspective, comparative studies on how the DataDiscovery tools are promoted on the market can be developed [12].

Conclusions

One of main objectivesfor the companies is to be competitve on the market. Delivering quality services and producing quality goods is not enough and companies have to adapt to the new economic climate and to the new customer's needs. This is not easy to do, but decisions can be made more easily with the help of the BI tools, especially, the Data Discovery tools. The classic reporting answers to predetermined-questions. Redports are built by the IT specialists and this might be time consuming, in some situations and offer useful, but in the same time limited and static information.Statistics address to advanced users, are difficult to customize and are powerful, but in the same time highly complex.

The Data Discovery tools offer real time answers to questions and enable faster and better decisions discovering new insights using the existing data. These tools can be used in various marketing fields such as: customer segmentation, planning and forecasting, supply chain management, building what-if scenarios, customer retention and churn analysis, supply chain optimization. Prtofolio management and retention policies.

The current article presented the data discovery tools, using mainly the demos available on the sites, from an end-user perspectives. Further case studies can be developed to see what are the results of implementing some of the decisions made after using such a tool.

Acknowledgement

This POSDRU/107/1.5/S/77213 „Ph.D. for a European standards" (DOCCENT).

career in interdisciplinary economic research at the

References

1. Berry M. J. A, Linoff G. S. (2011) Data Mining Techniques for Marketing, Sales and Customer Relationship Management - Third Edition, Wiley Publishig, p-1-86.

2. Mewada S , Singh UK (2012) Performance Analysis of Secure Wireless Mesh Network. RJRRS, 1(3):80-85

3. Tsipsis K, Choriannopoulos A (2009) Data Mining Techniques in CRM-Inside Customer Segmentation, Wiley Publishing, 17-127

4. Kolb J (2012) Business Intelligence in Plain Language: A practical guide to Data Mining and Business Analytics, Kindle Edition.

5. Mukhles Z (2009) Business Intelligence: Its Ins and Outs, Techn Eval Center.

6. Muntean M, Târnăveanu DA, (2010) BI Approach for Business Performance, 5th WSEAS Conference on Economy and Management Transformation.

7. Williams S , Williams N (2007) The Profit Impact of Business Intelligence, Morgan Kaufmann Publishing, pp 159-167.

8. Spotfire- http://www.screencast.com/t/mFFY0apGJ, 15 March 2013

9. Applied Data Lab http://www.applieddatalabs.com/content/data-discovery-tools-2012-roundup , 8 March 2013

10. Wikipedia http://en.wikipedia.org/wiki/QlikTech , 23 March 2013

11. Qlik View http://eu.demo.qlikview.com/ , 7 March 2013

12. Tableau software http://www.tableausoftware.com/, 7 March 2013